



caBIG

*cancer Biomedical
Informatics Grid*



caBIG Joint Architecture, V/CDE WS Meeting

Vocabulary Deployment Issues

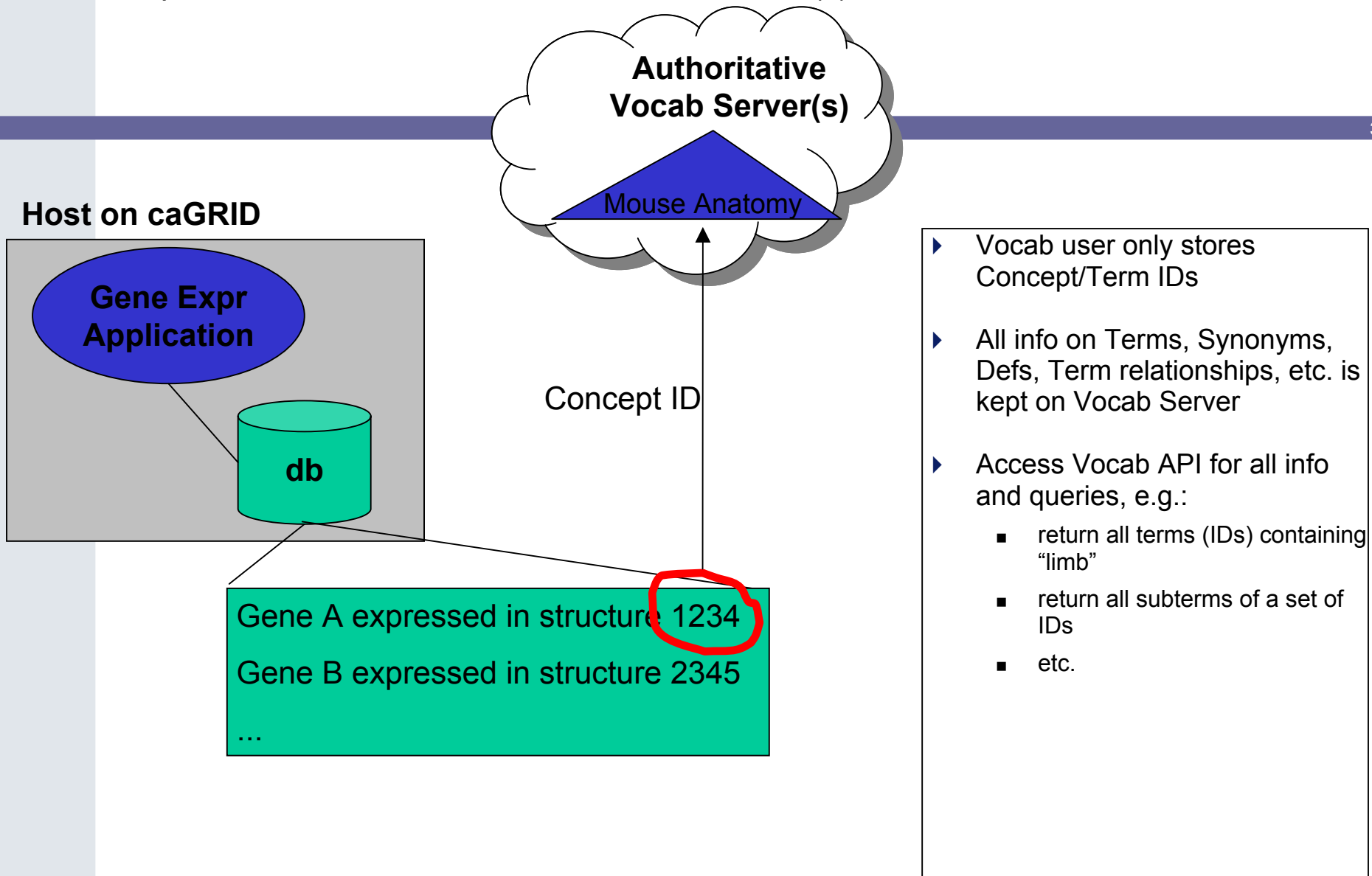
October 25-26, 2004

Jim Kadin

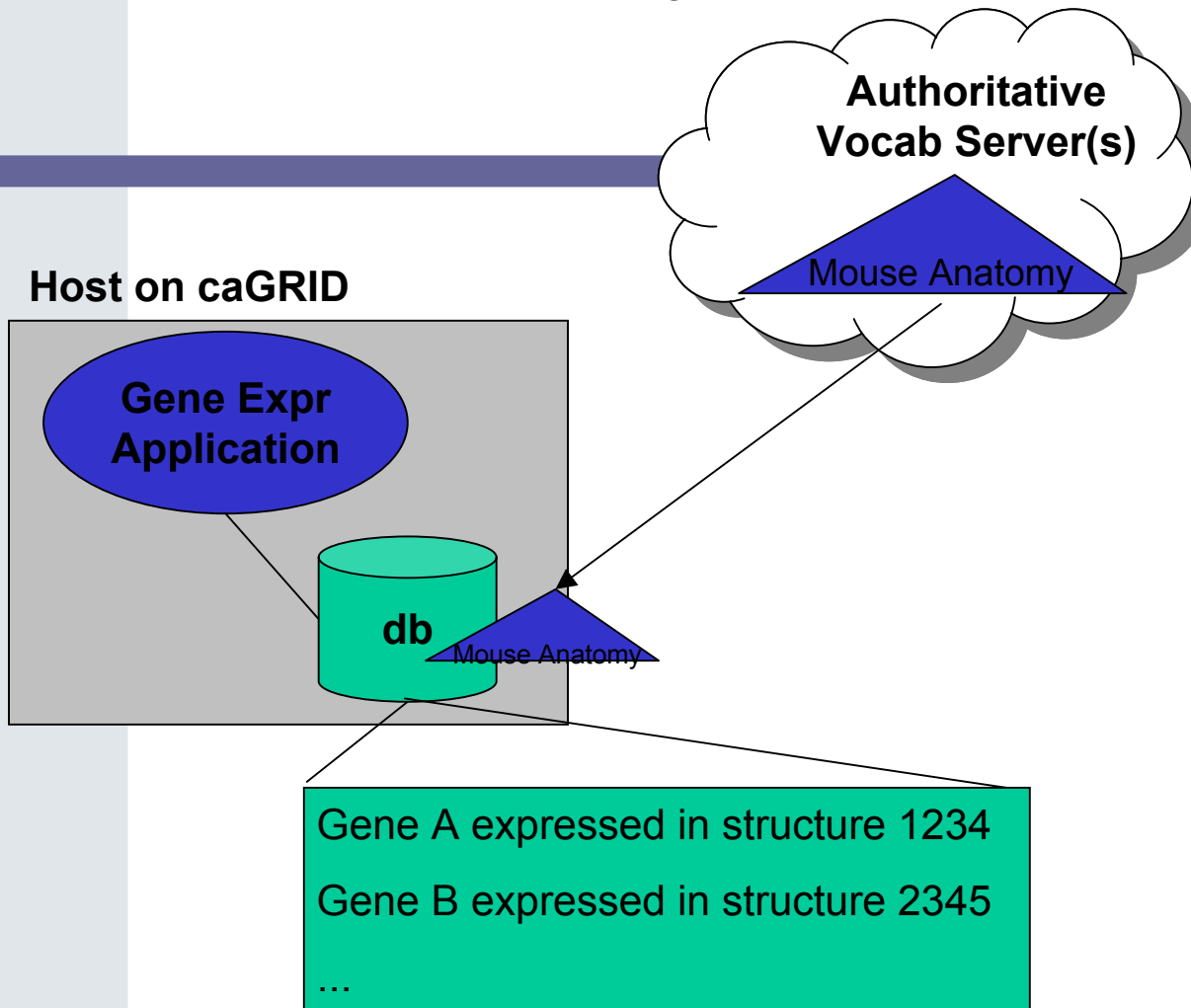
Frank Hartel

- ▶ How are caBIG vocabularies accessed and used by caBIG applications?
- ▶ Gene Expression, Mouse Anatomy Examples

Option 1: All Vocab Info Stored on Vocab Server(s)

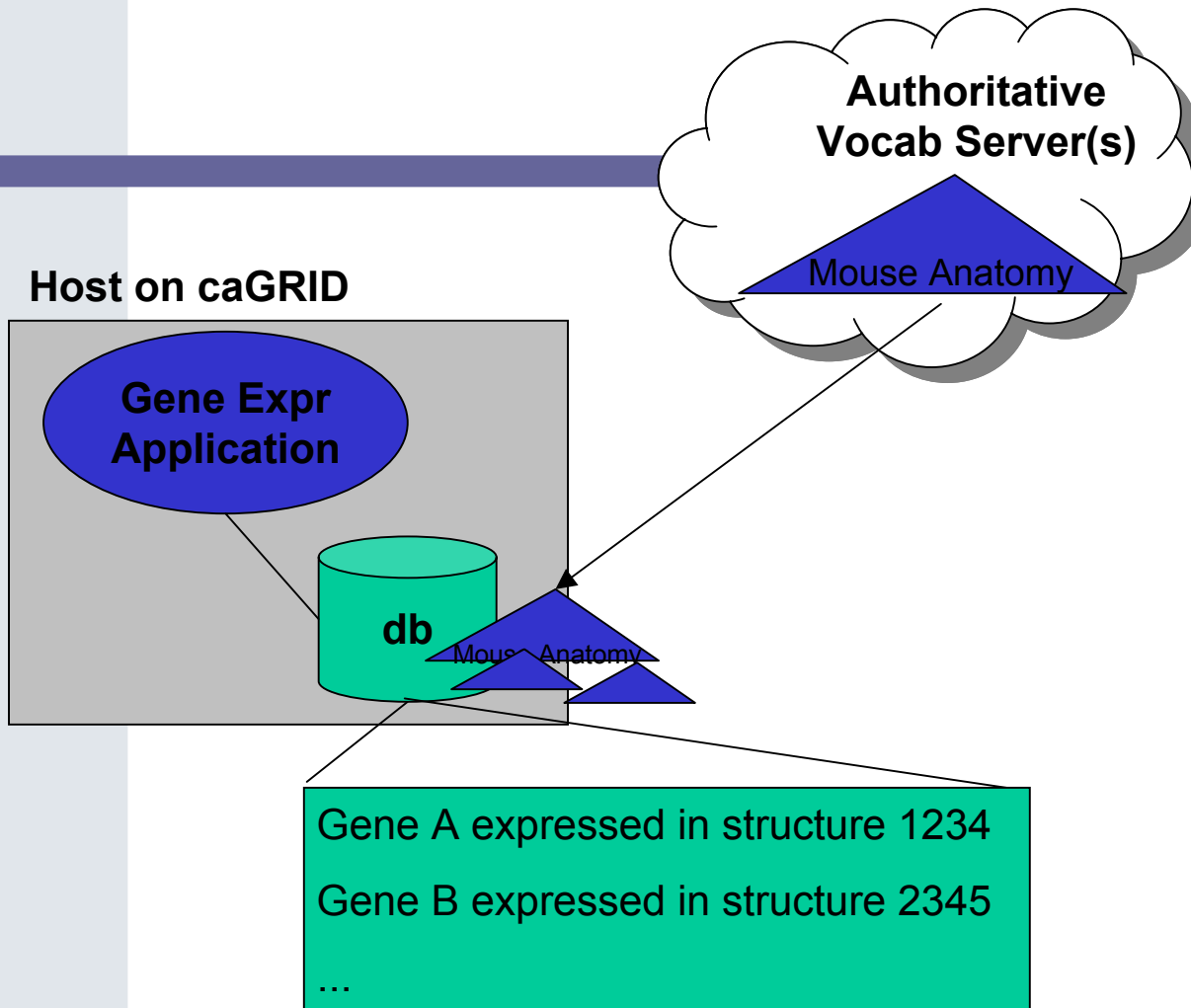


Option 2: Local Caching of Vocabs



- ▶ Cache for all the usual reasons
- ▶ Allows different data access patterns

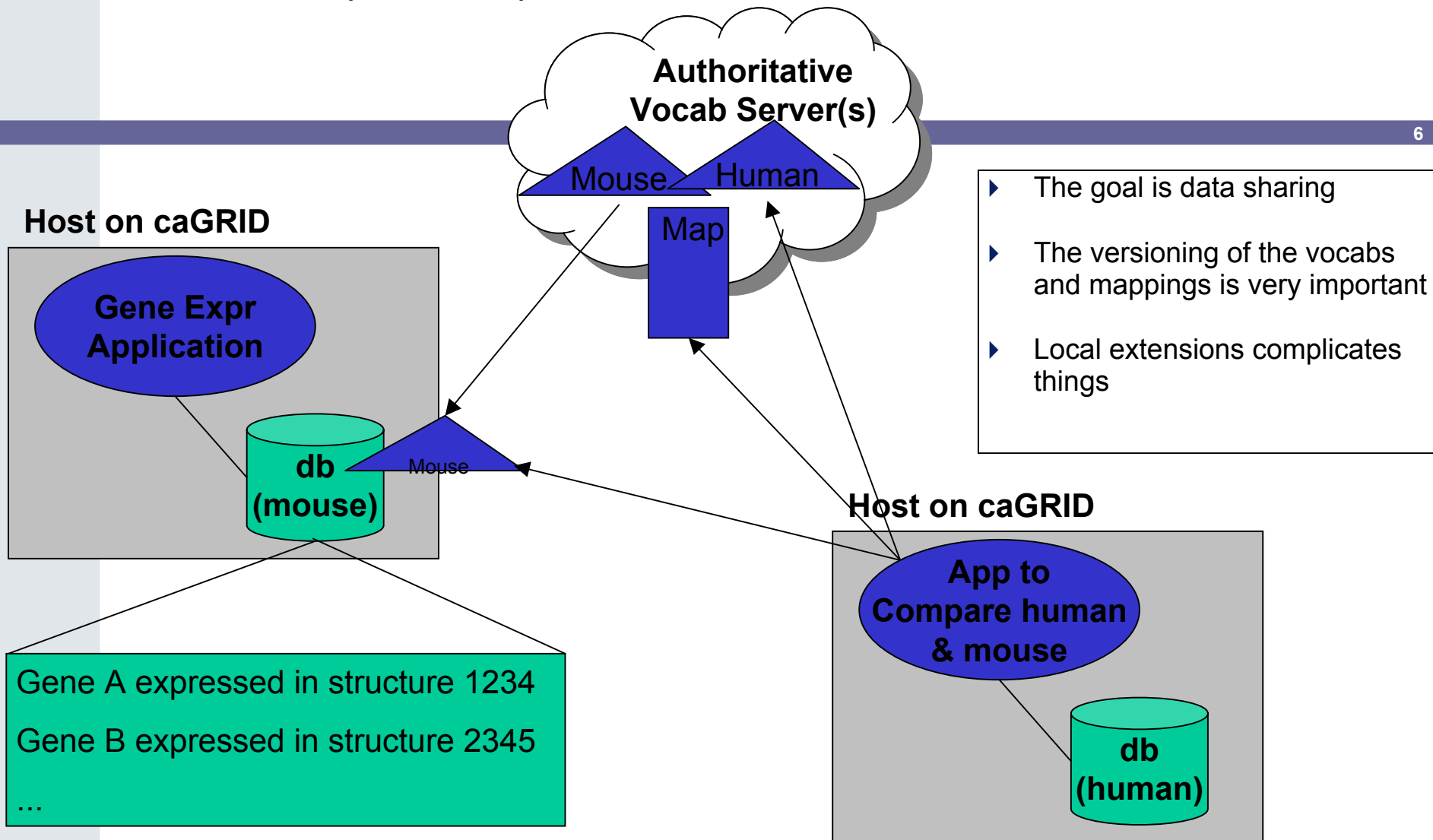
Option 2A: Local Extensions



- ▶ Allows site flexibility
- ▶ Complicates resynchronizing and other things

A More Complex Example

6



Issues, Questions, Constraints

7

- ▶ Is there one central vocab server or is it distributed? Serve local & global vocabs the same way?
- ▶ Need an ID for a vocab and a way to identify vocab versions
- ▶ Resynchronization/Vocab Versioning
- ▶ Local Extensions
- ▶ Concept IDs
- ▶ Mappings
- ▶ Tool support

Resynchronization, Vocab Versioning

8

- ▶ Vocab user (cached or uncached) needs to be able to refer to old versions of a vocab (i.e., decide their own resync schedule)
 - Vocab servers must maintain old vocab versions
 - Vocab users must keep track of their versions and proclaim their versions to other apps
- ▶ How does a site resync? How do we communicate changes?
 - changes to concept semantics (ID should change)
 - “trivial” changes – add synonyms, reword defs, ...
 - deletions, merges, splits
 - restructuring

Local Extensions

9

- ▶ Examples: adding concepts, restructuring, add/delete terms/synonyms
- ▶ Allows sites to get their work done w/o waiting
- ▶ Complicates resynchronization
- ▶ Do we support at all?
- ▶ Vocab ID or version for local extension?

Concept IDs

10

- ▶ For Imported Vocab, will caBIG assign its own ID to each concept or use externally assigned IDs? Maintain both?
- ▶ OIDs?
- ▶ LSIDs?

Mappings

11

- ▶ Example: mapping between human anatomy and mouse anatomy
- ▶ Is a mapping part of one vocabulary or an object in its own right?
- ▶ Vocab servers must provide mappings and mapping APIs
- ▶ Mapping IDs and versions

Tool Support

12

- ▶ Do we provide tools for cancer centers to build/maintain/access their own local vocabs?
 - data model
 - browser or browser toolkit
 - vocab editor or toolkit
- ▶ Do we provide tools for accessing vocabs at the central site?
 - browser or toolkit
 - vocab editor
 - resynchronization/caching tools